

## RATES OF STRONG CONVERGENCE FOR $U$ -STATISTICS IN FINITE POPULATIONS

P. N. KOKIC and N. C. WEBER

(Received 1 October 1989; revised 4 February 1990)

Communicated by Timothy C. Brown

### Abstract

Let  $U_{Nn}$  be a  $U$ -statistic based on a simple random sample of size  $n$  selected without replacement from a finite population of size  $N$ . Rates of convergence results in the strong law are obtained for  $U_{Nn}$ , which are similar to those known for classical  $U$ -statistics based on samples of independent and identically distributed (iid) random variables.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*) (1985 *Revision*): 60 F 15.

*Keywords and phrases*: rates of convergence, strong convergence,  $U$ -statistics, degenerate  $U$ -statistics, finite populations.

### 1. Introduction

$U$ -statistics, first introduced by Hoeffding (1948), form a large class of statistics that are widely used in estimation and testing. Nandi and Sen (1963) studied the weak limit behaviour of  $U$ -statistics based on samples drawn from finite populations, and Sen (1970) considered the strong convergence behaviour of such statistics.

Rates of convergence in the strong law of large numbers for  $U$ -statistics based on samples of independent and identically distributed (iid) random variables have been the focus of attention in several recent papers. Baum and Katz (1965) obtained necessary and sufficient conditions for convergence rate results for the mean of an iid sample: Lin (1981) obtained the corresponding sufficient condition for  $U$ -statistics, and Kokic (1987) extended these results

to obtain convergence rates for degenerate *U*-statistics. The aim of this paper is to extend these results to *U*-statistics based not on iid samples but on samples from finite populations.

Let  $\{x_{N1}, \dots, x_{NN}\}$  be a vector of  $N$  real numbers denoting the elements of the  $N$ th finite population of interest. Let  $(R_{N1}, \dots, R_{NN})$  be a vector chosen at random from the set of  $N!$  permutations of  $(1, 2, \dots, N)$  and let  $X_{Ni} = x_{NR_{Ni}}$ . A random sample of size  $n$  drawn from the finite population of size  $N$  can be written  $X_{N1}, \dots, X_{Nn}$ . An unbiased estimator of the population mean,  $\bar{x}_N = N^{-1} \sum_{i=1}^N x_{Ni}$ , is  $n^{-1}S_{Nn}$ , where  $S_{Nn} = \sum_{i=1}^n X_{Ni}$ . This is an example of a *U*-statistic of order 1. In this paper we will be mainly interested in *U*-statistics of order 2 defined as follows. Let  $h_N: \mathbb{R}^2 \rightarrow \mathbb{R}$  be a sequence of functions which are symmetric in their arguments and let  $w_{Nij} = h_N(x_{Ni}, x_{Nj})$ . If  $W_{Nij} = h_N(X_{Ni}, X_{Nj})$  is unbiased for some unknown population parameter  $\theta_N$ , say, then the corresponding *U*-statistic with kernel  $h_N$ ,  $\binom{n}{2}^{-1}U_{Nn}$ , is also unbiased for  $\theta_N$ , where

$$U_{Nn} = \sum_{i=2}^n \sum_{j=1}^{i-1} W_{Nij}.$$

Note that if we take  $h_N(x_{Ni}, x_{Nj}) = \frac{1}{2}(x_{Ni} - x_{Nj})^2$  then we obtain the usual estimator of the population variance. This paper will study the conditions under which  $\{\binom{n}{2}^{-1}U_{Nn} - \theta_N\}$  converges strongly to zero and the rate of such convergence. Strong convergence for such situations was defined by Sen (1970). An equivalent definition of strong convergence is given below.

**DEFINITION.** Let  $\{T_{Nn}(X_{N1}, \dots, X_{Nn}), 1 \leq n \leq N, N \geq 1\}$  be a triangular array of statistics and let  $\{\theta_N, N \geq 1\}$  be a sequence of real numbers. Then  $T_{Nn} - \theta_N$  strongly converges (s.c.) to 0, if for every  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} |T_{Nk} - \theta_N| > \varepsilon \right) = 0.$$

Rosén (1964) established that  $n^{-1}S_{Nn} - \bar{x}_N \rightarrow 0$  s.c. if and only if

$$(1.1) \quad \lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N |x_{Ni} - \bar{x}_N| I(|x_{Ni} - \bar{x}_N| > L) = 0.$$

In Section 2 we will show that the rate at which  $n^{-1}S_{Nn} - \bar{x}_N$  converges to zero depends on the values of  $t$  for which  $\sup_{N \geq 1} N^{-1} \sum_{i=1}^N |x_{Ni} - \bar{x}_N|^t < \infty$ . In Section 3 we establish the corresponding result for  $U_{Nn}$ , and improve the strong convergence result for  $U_{Nn}$  established in Theorem 2 of Sen (1970). Further, we refine the rate of convergence for the case when the arrays  $\{w_{Nij}\}$  have all row and column sums equal to a fixed constant. These results are

analogous to the rate results established by Kokic (1987) for degenerate  $U$ -statistics based on samples of iid random variables.

**2. Rate results for the sample sum**

Suppose

$$(2.1) \quad \sup_{N \geq 1} N^{-1} \sum_{i=1}^N |x_{Ni}|^t < \infty.$$

Then (1.1) holds if  $t > 1$ . Hence  $n^{-1}S_{Nn} - \bar{x}_N \rightarrow 0$  s.c.. In fact, we have the following result.

**THEOREM 1.** *If  $0 < t < r$ ,  $r > 1$  and condition (2.1) holds, then for each  $\varepsilon > 0$ ,*

$$(2.2) \quad \limsup_{N \rightarrow \infty} \sum_{n=1}^N n^{r-2} P \left( \max_{n \leq k \leq N} k^{-r/t} |S_{Nk}| > \varepsilon \right) < \infty.$$

*If  $1 < r \leq t \leq 2r$ , condition (2.1) holds and for all  $N \geq 1$ ,  $\bar{x}_N = 0$ , then for each  $\varepsilon > 0$ , (2.2) holds.*

This result is the analogue of Theorem 3 of Baum and Katz (1965), who developed a similar result for the sum of iid random variables. The following lemma about triangular arrays of real numbers allows us to deduce the convergence rate result given in Corollary 1 of Theorem 1.

**LEMMA 1.** *Let  $\{a_{Nn}, 1 \leq n \leq N, N \geq 1\}$  be a triangular array of positive numbers satisfying  $a_{Nn} \geq a_{N, n+1}$  when  $N > n \geq 1$ . If for some  $r > 1$ ,*

$$\limsup_{N \rightarrow \infty} \sum_{n=1}^N n^{r-2} a_{Nn} < \infty$$

then

$$\limsup_{N \rightarrow \infty} a_{Nn} = O(n^{-r+1}) \text{ as } n \rightarrow \infty.$$

**COROLLARY 1.** *Under the conditions of Theorem 1*

$$\limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} k^{-r/t} |S_{Nk}| > \varepsilon \right) = O(n^{-r+1}) \text{ as } n \rightarrow \infty.$$

The large order term in Corollary 1 may be replaced by the corresponding small order term if we replace (2.1) by the stronger condition

$$(2.3) \quad \lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N |x_{Ni}|^t I(|x_{Ni}| > L) = 0.$$

**THEOREM 2.** *Under the conditions of Theorem 1, if (2.3) hold then*

$$\limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} k^{-r/t} |S_{Nk}| > \varepsilon \right) = o(n^{-r+1}) \quad \text{as } n \rightarrow \infty.$$

Suppose  $\bar{x}_N = 0$ . Condition (1.1) is equivalent to the sequence of random variables  $\{X_{N1}, N \geq 1\}$  being uniformly integrable, which is equivalent to  $n^{-1}S_{Nn} \rightarrow 0$  s.c. For  $t > 1$  condition (2.1) implies the uniform integrability of  $\{X_{N1}, N \geq 1\}$ , and futher implies the range of convergence rate results given in Corollary 1. Condition (2.3) is equivalent to the uniform integrability of  $\{|X_{N1}|^t, N \geq 1\}$ . It implies condition (2.1) and the slightly stronger result of Theorem 2.

### 3. Rate results for *U*-statistics

Throughout this section it is assumed that

$$(3.1) \quad \sum_{i=2}^N \sum_{j=1}^{i-1} w_{Nij} = 0.$$

Also, by symmetry, and since the diagonal terms may be defined arbitrarily, we have

$$(3.2) \quad w_{Nij} = w_{Nji} \quad \text{and} \quad w_{Nii} = 0.$$

Sen (1970) showed that  $\binom{n}{2}^{-1}U_{Nn} \rightarrow 0$  s.c. if  $\sup_{N > 2} \binom{N}{2}^{-1} \sum_{i=2}^N \sum_{j=1}^{i-1} w_{Nij}^2 < \infty$ . In Theorem 3 below the same result is established under a weaker set of conditions. Let

$$\hat{w}_{Ni} = (N - 2)^{-1} \sum_{j=1}^N w_{Nij}$$

and

$$(3.3) \quad v_{Nij} = \begin{cases} w_{Nij} - \hat{w}_{Ni} - \hat{w}_{Nj} & \text{if } i \neq j, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

**THEOREM 3.** *Suppose that*

$$\lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N |\hat{w}_{Ni}| I(|\hat{w}_{Ni}| > L) = 0$$

and

$$\lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \binom{N}{2}^{-1} \sum_{i=2}^N \sum_{j=1}^{i-1} |v_{Nij}| I(|v_{Nij}| > L) = 0.$$

Then  $\binom{n}{2}^{-1} U_{Nn} \rightarrow 0$  s.c.

Corresponding to Theorem 1 we have the convergence rate result given in Theorem 4 below. The result is essentially the analogue of Theorem 2 of Lin (1981). However, we first give the following moment bound result, which is essential in the proof of Theorems 4 and 5. Let  $\nu_{Nt} = \binom{N}{2}^{-1} \sum_{i=2}^N \sum_{j=1}^{i-1} |w_{Nij}|^t$ .

**LEMMA 2.** *Suppose that  $U_{Nn}$  is based on an array  $\{w_{Nij}\}$  satisfying (3.1), (3.2) and for all  $i$ ,  $\sum_{j=1}^N w_{Nij} = 0$ . Then for each  $t \geq 2$  there exists a constant  $K_t$ , such that*

$$E|U_{Nn}|^t \leq K_t(Npq)^t \nu_{Nt}.$$

**THEOREM 4.** *Suppose  $r > 1$  and that for some  $2 \leq t < 2r$ ,*

$$(3.4) \quad \sup_{N \geq 1} \nu_{Nt} < \infty.$$

Then for each  $\varepsilon > 0$ ,

$$\limsup_{N \rightarrow \infty} \sum_{n=2}^N n^{r-2} P\left(\max_{n \leq k \leq N} k^{-1-r/t} |U_{Nk}| > \varepsilon\right) < \infty.$$

**COROLLARY 2.** *Under the conditions of Theorem 4*

$$\limsup_{N \rightarrow \infty} P\left(\max_{n \leq k \leq N} k^{-1-r/t} |U_{Nk}| > \varepsilon\right) = O(n^{-r+1}) \text{ as } n \rightarrow \infty.$$

If in addition to (3.4) we assume that

$$(3.5) \quad \lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N |\hat{w}_{Ni}|^t I(|\hat{w}_{Ni}| > L) = 0,$$

then the large order term in Corollary 2 may be replaced by the corresponding small order term.

**THEOREM 5.** *If the conditions Theorem 4 are satisfied and (3.5) holds, then for each  $\epsilon > 0$*

$$\limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} k^{-1-r/t} |U_{Nk}| > \epsilon \right) = o(n^{-r+1})$$

as  $n \rightarrow \infty$ .

If  $U_{Nn}$  is based on an array  $\{w_{Nij}\}$  with the property that the row and column sums of  $\{w_{Nij}\}$  all equal the same constant, then we call  $U_{Nn}$  a degenerate *U*-statistic. Condition (3.1) forces a degenerate *U*-statistic to satisfy

$$(3.6) \quad \sum_{i=1}^N w_{Nij} = \sum_{j=1}^N w_{Nij} = 0.$$

The reason we have defined  $\{v_{Nij}\}$  as at (3.3) is that it satisfies (3.6) in addition to (3.1) and (3.2). As shown in Lemma 2, it is possible to obtain high order moment bounds for degenerate *U*-statistics, which are crucial in obtaining convergence rate results for *U*-statistics. The following strong convergence rate result for degenerate *U*-statistics is the analogue of Theorem 1 of Kocic (1987).

**THEOREM 6.** *Suppose  $\{U_{Nn}\}$  are degenerate *U*-statistics satisfying (3.1) and (3.2). Let  $r$  be a positive constant and suppose that for some  $t \geq 2$ , condition (3.4) is satisfied. Then for any  $\delta > 1$  and any  $\epsilon > 0$ ,*

$$\limsup_{N \rightarrow \infty} \sum_{n=2}^N n^{r-\delta} P \left( \max_{n \leq k \leq N} k^{-1-r/t} |U_{Nk}| > \epsilon \right) < \infty,$$

and

$$\limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} k^{-1-r/t} |U_{Nk}| > \epsilon \right) = O(n^{-r}).$$

We see, on comparing this result with Theorem 4 and Corollary 2, that the rate at which  $\binom{n}{2}^{-1} U_{Nn}$  converges to zero (s. c.) is faster for degenerate than nondegenerate *U*-statistics.

### 4. Proofs

Throughout let  $c_1, c_2, \dots$  denote positive constants which do not depend on  $n$  or  $N$ . The proof of Theorem 1 is reduced to checking that three sums converge. The following technical lemma is required to bound the third of these sums. The  $\gamma$  parameter introduced in the lemma will be used to

determine truncation values in the proof of the theorem. This truncation technique can also be used to correct the proof of the claim in Baum and Katz (1965), page 109, that given  $E|X|^t < \infty$  for some  $0 < t < 1$  then  $n^{1-1/t}|EXI(|X| < n^{1/t})| \rightarrow 0$  as  $n \rightarrow \infty$ . To see this note that for any  $\gamma \in (0, 1)$ ,

$$\begin{aligned} & n^{1-1/t}|EXI(|X| < n^{1/t})| \\ & \leq n^{1-1/t}\{E|X|I(|X| \leq n^{\gamma/t}) + E|X|I(n^{\gamma/t} < |X| < n^{1/t})\} \\ & \leq n^{(1-\gamma)(1-1/t)}E|X|^t + E(|X|^t I(n^\gamma < |X|^\gamma)) \\ & \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

**LEMMA 3.** *Suppose that  $r > 1$  and  $2r > t > 0$ . Let  $\gamma$  be any number satisfying  $1 > \gamma > \max\{r^{-1}, t/(2r)\}$ . Let  $y_{N1}, \dots, y_{NN}$  be real numbers such that for some  $c_1$  and  $i = 1, \dots, N$ ,  $|y_{Ni}| < c_1 n^{\gamma r/t}$ , where  $1 \leq n \leq N$ ,  $\sum_{k=1}^N y_{Nk} = 0$  and  $\sup_{N \geq 1} N^{-1} \sum_{k=1}^N |y_{Nk}|^t < \infty$ . Let  $Y_{Ni} = y_{NR_{Ni}}$  and let  $j$  be the smallest integer greater than or equal to  $t$ . For each integer  $M$ , independent of  $n$  and satisfying*

$$(4.1) \quad 2Mj\gamma r/t - \gamma r + 1 \geq Mj$$

there exists a  $c_2$  such that

$$E \left| \sum_{k=1}^n Y_{Nk} \right|^{2Mj} \leq c_2 n^{(2Mj\gamma r/t - \gamma r + 1)}.$$

**PROOF.** The key to establishing this lemma is the useful inequality found in Theorem 4 of Hoeffding (1963), which states

$$(4.2) \quad E \left| \sum_{k=1}^n Y_{Nk} \right|^{2Mj} \leq E \left| \sum_{k=1}^n Z_{Nk} \right|^{2Mj},$$

where  $Z_{N1}, \dots, Z_{Nn}$  are iid random variables uniformly distributed on  $y_{N1}, \dots, y_{NN}$ . The result now follows from (4.2) by the same argument as used by Katz (1963) to bound his expression (7).

**PROOF OF THEOREM 1.** Without loss of generality assume that  $N \geq 2$  and  $\varepsilon = 2$ . For convenience of notation we will drop the  $N$  subscript in the double subscripts throughout all proofs, except where extra clarity is required. Choose a  $\gamma$  and an integer  $M$  so that (4.1) is satisfied and

$$(4.3) \quad (Mt - 1)/(Mt) > \gamma > \max\{(r + 1)/(2r), t/(2r)\}.$$

For  $1 \leq m \leq N$  let

$$A_m^{(1)} = \bigcup_{k=1}^m \{|X_k| > m^{r/t}\}, \quad A_m^{(2)} = \bigcup_{k_2=2}^m \bigcup_{k_1=1}^{k_2-1} \{|X_{k_1}|, |X_{k_2}| > m^{r/t}\}$$

and

$$A_m^{(3)} = \left\{ \left| \sum_{k=1}^m X_{km} \right| > m^{r/t} \right\},$$

where  $X_{km} = X_k$  if  $|X_k| < m^{r/t}$ , and  $X_{km} = 0$  otherwise. If we argue as in (1) of Katz (1963), to prove the theorem it suffices to show that

$$(4.4) \quad \limsup_{N \rightarrow \infty} \sum_{n=1}^N n^{r-2} P \left( \bigcup_{m=n}^N A_m^{(l)} \right) < \infty \quad \text{for } l = 1, 2 \text{ and } 3.$$

We begin by proving (4.4) for  $l = 1$ . Note that

$$(4.5) \quad \bigcup_{m=n}^N A_m^{(1)} = \bigcup_{k=1}^n \{|X_k| > n^{r/t}\} \cup \bigcup_{k=n+1}^N \{|X_k| > k^{r/t}\},$$

so  $\sum_{n=1}^N n^{r-2} P(\bigcup_{m=n}^N A_m^{(1)})$  can be expressed as the sum of two terms each of which can be bounded by

$$(4.6) \quad \sum_{n=1}^N n^{r-1} P(|X_1| > n^{r/t}) \leq c_3 E|X_1|^t.$$

Thus, by condition (2.1), (4.4) follows in the case  $l = 1$ .

Now for  $l = 2$

$$\bigcup_{m=n}^N A_m^{(2)} = \bigcup_{k_1=2}^n \bigcup_{k_2=1}^{k_1-1} \{|X_{k_1}|, |X_{k_2}| > n^{r/t}\} \cup \bigcup_{k_1=n+1}^N \bigcup_{k_2=1}^{k_1-1} \{|X_{k_1}|, |X_{k_2}| > k_1^{r/t}\}$$

and so by similar techniques to those used in the case  $l = 1$ ,

$$\sum_{n=1}^N n^{r-2} P \left( \bigcup_{m=n}^N A_m^{(2)} \right) \leq c_4 \sum_{k=1}^N k^r P(|X_1|, |X_2| > k^{r/t}).$$

Note  $X_1$  and  $X_2$  are not independent but

$$\begin{aligned} P(|X_1|, |X_2| > k^{r/t}) &= [N(N-1)]^{-1} \sum_{k_1 \neq k_2} I(|x_{k_1}| > k^{r/t}) I(|x_{k_2}| > k^{r/t}) \\ &\leq N/(N-1) \left\{ N^{-1} \sum_{k_1=1}^N I(|x_{k_1}| > k^{r/t}) \right\}^2 \\ &\leq 2\{E|X_1|^t/k^{r/t}\}^2. \end{aligned}$$



Thus

$$(4.7) \quad \sum_{n=1}^N n^{r-2} P \left( \bigcup_{m=n}^N A_m^{(2)} \right) \leq c_5 (E|X_1|^t)^2 \sum_{k=1}^{\infty} k^{r-2\gamma r}.$$

By (4.3),  $r(1 - 2\gamma) < -1$ , so the series at (4.7) converges. Thus by (2.1), (4.4) follows in the case  $l = 2$ .

It remains to prove (4.4) in the case  $l = 3$ . Let  $Y_{km} = X_{km} - EX_{km}$ ; then

$$(4.8) \quad A_m^{(3)} \subset \left\{ \left| \sum_{k=1}^m Y_{km} \right| > m^{r/t} (1 - m^{1-r/t} |EX_{1m}|) \right\}.$$

We now show for sufficiently large  $m$  that  $m^{1-r/t} |EX_{1m}|$  is small.

When  $1 < r \leq t$ ,  $EX_k = 0$ , and so

$$\begin{aligned} m^{1-r/t} |EX_{1m}| &\leq m^{1-r/t} E\{|X_1|(|X_1|m^{-\gamma r/t})^{t-1} I(|X_1| \geq m^{\gamma r/t})\} \\ &\leq E|X_1|^t m^{(1-\gamma r)-r(1-\gamma)/t}, \end{aligned}$$

and when  $0 < t < r$  and  $r > 1$ ,

$$m^{1-r/t} |EX_{1m}| \leq m^{(1-\gamma)(1-r/t)} E|X_1|^{t/r}.$$

Now (4.3) ensures that  $(1-\gamma r)-r(1-\gamma)/t < 0$  and  $(1-\gamma)(1-r/t) < 0$ , and so condition (2.1) ensures for all sufficiently large  $m$  that  $m^{1-r/t} |EX_{1m}| < \frac{1}{2}$ . By (4.8), to complete the proof it suffices to show that

$$(4.9) \quad \limsup_{N \rightarrow \infty} \sum_{n=1}^N n^{r-2} P \left( \bigcup_{m=n}^N \left\{ \left| \sum_{k=1}^m Y_{km} \right| > \frac{1}{2} m^{r/t} \right\} \right) < \infty.$$

The sum on the left-hand side of (4.9) is bounded by a constant multiple of

$$\begin{aligned} &\sum_{m=1}^N m^{r-1} P \left( \left| \sum_{k=1}^m Y_{km} \right| > \frac{1}{2} m^{r/t} \right) \\ &\leq c_6 \sum_{m=1}^N m^{r-1} m^{-2Mjr/t} E \left| \sum_{k=1}^m Y_{km} \right|^{2Mj} \\ (4.10) \quad &\leq c_7 \sum_{m=1}^{\infty} m^{(2Mjr/t)(\gamma-1)-r(\gamma-1)}, \end{aligned}$$

by Lemma 3. Now (4.1) and (4.3) ensure that the series at (4.10) converges, and hence (4.9) and Theorem 1 follow.

Lemma 1 may be proved with standard analytic techniques. The details are therefore omitted from this paper.

**PROOF OF THEOREM 2.** With the notation of the proof of Theorem 1 it is sufficient to show that

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} n^{r-1} P \left( \bigcup_{m=n}^N A_m^{(l)} \right) = 0 \quad \text{for } l = 1, 2 \text{ and } 3.$$

The argument proceeds almost exactly as in Theorem 1, where we use (2.3) to bound the  $l = 1$  terms, that is, to bound  $n^{r-1} P(\bigcup_{k=1}^n \{X_k > n^{r/t}\})$  and  $n^{r-1} P(\bigcup_{k=n+1}^N \{X_k > k^{r/t}\})$  by  $E|X_1|^t I(|X_1|^t > n^{r/t})$ .

Essential to the proof of Theorems 3, 4 and 5 is the following decomposition. Note that

$$(4.11) \quad \binom{n}{2}^{-1} U_{Nn} = 2n^{-1} \widehat{S}_{Nn} + \binom{n}{2}^{-1} V_{Nn},$$

where

$$V_{Nn} = \sum_{l=2}^n \sum_{j=1}^{l-1} v_{NR_{Ni}R_{Nj}} \quad \text{and} \quad \widehat{S}_{Nn} = \sum_{i=1}^n \widehat{w}_{NR_{Ni}}.$$

**PROOF OF THEOREM 3.** Assume without loss of generality that  $N \geq 3$ . By (4.11),

$$(4.12) \quad P \left( \max_{n \leq k \leq N} \binom{k}{2}^{-1} |U_k| > \varepsilon \right) \leq P \left( \max_{n \leq k \leq N} 2k^{-1} |\widehat{S}_k| > \frac{1}{2} \varepsilon \right) + P \left( \max_{n \leq k \leq N} \binom{k}{2}^{-1} |V_k| > \frac{1}{2} \varepsilon \right).$$

By Theorem 5.2 of Rosén (1964),

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} 2k^{-1} |\widehat{S}_k| > \varepsilon \right) = 0.$$

To bound the second term we note that by Lemma 2.1 of Sen (1970),  $\{\binom{n}{2}^{-1} V_{Nn}, 2 \leq n \leq N\}$  is a reverse martingale. Thus,

$$P \left( \max_{n \leq k \leq N} \binom{k}{2}^{-1} |V_k| > \frac{1}{2} \varepsilon \right) \leq 2\varepsilon^{-1} \binom{n}{2}^{-1} E|V_n|.$$

To bound  $E|V_n|$  we let  $v_{ij} = a_{ij} + b_{ij}$ , where

$$\begin{aligned} b_{ij} &= v_{ij} I(|v_{ij}| > L) - (N-2)^{-1} \sum_{j=1}^N v_{ij} i I(|v_{ij}| > L) \\ &\quad - (N-2)^{-1} \sum_{i=1}^N v_{ij} I(|v_{ij}| > L) \\ &\quad + (N-1)^{-1} (N-2)^{-1} \sum_{i=1}^N \sum_{j=1}^N v_{ij} I(|v_{ij}| > L), \quad \text{if } i \neq j \end{aligned}$$

and

$$b_{ij} = 0, \text{ if } i = j.$$

Thus  $\sum_{i=1}^N a_{ij} = \sum_{i=1}^N b_{ij} = 0$ , and for some  $c_\delta$ ,

$$\begin{aligned}
E|V_n| &\leq \sum_{i=2}^n \sum_{j=1}^{i-1} E|b_{R_i R_j}| + \left\{ E \left( \sum_{i=2}^n \sum_{j=1}^{i-1} a_{R_i R_j} \right)^2 \right\}^{1/2} \\
&\leq \binom{n}{2} \binom{N}{2}^{-1} \sum_{i=2}^n \sum_{j=1}^{i-1} |b_{ij}| + \left\{ c_8 \binom{n}{2} \binom{N}{2}^{-1} \sum_{i=2}^n \sum_{j=1}^{i-1} a_{ij}^2 \right\}^{1/2}.
\end{aligned}$$

Since  $\binom{N}{2}^{-1} \sum_{i=2}^N \sum_{j=1}^{i-1} a_{ij}^2$  is bounded by a constant,

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left( \max_{n \leq k \leq N} \binom{k}{2}^{-1} |V_{Nk}| > \frac{1}{2}\epsilon \right) = 0.$$

This completes of the proof of Theorem 3.

**PROOF OF LEMMA 2.** Since  $\{\binom{N-n+2}{2} U_{N-n+2}, 2 \leq n \leq N\}$  is a martingale, we have via the theorem in Dharmadhikari et al. (1968) that

$$(4.13) \quad E \left| \binom{n}{2}^{-1} U_n \right|^t \leq c_9 (N-n)^{t/2} \max_{n \leq j \leq N} E \left| \binom{j}{2}^{-1} U_j - \binom{j+1}{2}^{-1} U_{j+1} \right|^t$$

where  $c_9 = \max\{8(t-1) \max(1, 2^{t-3})\}^t$ . After some algebra, it may be shown that

$$\begin{aligned}
&\binom{n+1}{2}^{-1} U_{n+1} - \binom{n}{2}^{-1} U_n \\
&= (n-1)(n+1)^{-1} \binom{n}{2}^{-1} \sum_{j=1}^n W_{n+1,j} - (n+1)^{-1} \binom{n}{2}^{-1} \sum_{i \neq j}^n \sum W_{ij}.
\end{aligned}$$

Hence by Hölder’s inequality,

$$\begin{aligned}
(4.14) \quad &E \left| \binom{n}{2}^{-1} U_n - \binom{n+1}{2}^{-1} U_{n+1} \right|^t \\
&\leq 2^{t-1} \binom{n}{2}^{-t} \left\{ E \left| \sum_{j=1}^n W_{n+1,j} \right|^t + E \left| \sum_{j=1}^{n-1} W_{nj} \right|^t \right\}.
\end{aligned}$$

Conditional on  $R_n = m$ ,  $\sum_{j=1}^{n-1} W_{nj}$  is the sum of a simple random sample of size  $n - 1$  selected without replacement from

$$\mathcal{P} = \{w_{m1}, \dots, w_{m,m-1}, w_{m,m+1}, \dots, w_{mN}\}.$$

Since  $U_n$  is degenerate the elements of  $\mathcal{P}$  sum to zero. Thus by Theorem 4 of Hoeffding (1963), and the theorem in Dharmadhikari et al. (1968),

$$E \left| \sum_{j=1}^{n-1} W_{nj} \right|^t \leq c_{10} (n-1)^{t/2} \nu_{nt}.$$

Since

$$(4.15) \quad \sum_{j=1}^{n-1} W_{nj} = - \sum_{j=n+1}^N W_{nj} \stackrel{\mathcal{Q}}{=} - \sum_{j=1}^{N-n} W_{N-n+1, j},$$

we have

$$(4.16) \quad E \left| \sum_{j=1}^{n-1} W_{nj} \right|^t \leq c_{10} (N-n)^{t/2} \nu_{nt},$$

and so by (4.13), (4.14) and (4.16),

$$(4.17) \quad E|U_n|^t \leq c_{10}^2 2^t (N-n)^t \nu_{Nt}.$$

By similar reasoning to that used at (4.15), we have  $U_n \stackrel{\mathcal{Q}}{=} U_{N-n}$ . The lemma follows from this fact and expression (4.17).

**PROOF OF THEOREM 4.** Using (4.12), we have

$$(4.18) \quad P \left( \max_{n \leq k \leq N} k^{-1-r/t} |U_k| > \varepsilon \right) \leq P \left( \max_{n \leq k \leq N} k^{-r/t} |\widehat{S}_k| > \frac{1}{2}\varepsilon \right) + P \left( \max_{n \leq k \leq N} k^{1-r/t} \left( \binom{k}{2} \right)^{-1} |V_k| > \varepsilon \right).$$

From Theorem 1,

$$\limsup_{N \rightarrow \infty} \sum_{n=2}^N n^{r-2} P \left( \max_{n \leq k \leq N} k^{-r/t} |\widehat{S}_k| > \frac{1}{2}\varepsilon \right) < \infty,$$

so the result will follow if

$$\limsup_{N \rightarrow \infty} \sum_{n=2}^N n^{r-2} P \left( \max_{n \leq k \leq N} k^{1-r/t} \left| \left( \binom{k}{2} \right)^{-1} V_k \right| > \varepsilon \right) < \infty.$$

Since  $\{ \left( \binom{k}{2} \right)^{-1} V_k, 2 \leq k \leq N \}$  is a reverse martingale, we have

$$(4.19) \quad P \left( \max_{n \leq k \leq N} k^{1-r/t} \left| \left( \binom{k}{2} \right)^{-1} V_k \right| > \varepsilon \right) \leq \binom{n}{2}^{-t} E|V_n|^t \varepsilon^{-t} n^{t-r},$$

and the result follows if we use the bound in Lemma 2 and argue as in Lin (1981), Theorem 2.1

Theorem 5 now follows almost immediately if we use Theorem 2 to deal with the first term in the bound (4.18), and (4.19) to bound the second term.

Theorem 6 is established by our using similar arguments to those used in Theorem 1 of Kokic (1987) and so the details are not included here.

### Acknowledgement

The authors wish to thank the referee for his helpful comments.

### References

- L. E. Baum and M. Katz (1965), 'Convergence rates in the law of large numbers', *Trans. Amer. Math. Soc.* **120**, 108–123.
- S. W. Dharmadhikari, V. Fabian and K. Jogdeo (1968), 'Bounds on the moments of martingales', *Ann. Math. Statist.* **39**, 1719–1723.
- W. Hoeffding (1948), 'A class of statistics with asymptotically normal distribution', *Ann. Math. Statist.* **19**, 293–325.
- W. Hoeffding (1963), 'Probability inequalities for sums of bounded random variables', *J. Amer. Statist. Assoc.* **58**, 13–30.
- M. Katz (1963), 'The probability in the tail of a distribution', *Ann. Math. Statist.* **34**, 312–318.
- P. N. Kokic (1987), 'Rates of convergence in the strong law of large numbers for degenerate  $U$ -statistics', *Statist. Probab. Letters* **5**, 371–374.
- K. Lin (1981), 'Convergence rate and the first exit time for  $U$ -statistics', *Bull. Inst. Math. Acad. Sin.* **9**, 129–143.
- H. K. Nandi and P. K. Sen (1963), 'Unbiased estimation of the parameters of a finite population', *Calcutta Statist. Assn. Bull.* **12**, 124–148.
- B. Rosén (1964), 'Limit theorems for sampling from a finite population', *Arch. Mat.* **5**, 383–424.
- P. K. Sen (1970), 'The Hájek-Rényi inequality for sampling from a finite population', *Sankhya A* **32**, 181–188.

Department of Mathematical Statistics  
University of Sydney  
N.S.W. 2006  
Australia